The Sound of Midrange Horns for Studio Monitors*

KEITH R. HOLLAND, *AES Member*, **and FRANK J. FAHY,** *AES Member*

Institute of Sound and Vibration Research, University of Southampton, Southampton SO17 1BJ, UK

AND

PHILIP R. NEWELL, *AES Member*

Vigo, Spain

A blind listening test is described in which 16 loudspeakers are compared with four reference loudspeakers under anechoic conditions. The test is concerned with the perceived sonic similarity between midrange horn loudspeakers and direct radiators and is intended to pinpoint the physical cause of a "characteristic sound" attributed to many studio monitor systems equipped with midfrequency-range horns. Comparisons are made between the listening test results and measurements of on-axis frequency response. The results indicate that short horns sound more similar to direct-radiating loudspeakers than long horns. It is concluded that the reflections from the mouth termination of long horns is responsible for the characteristic sound and that for studio monitor applications, a midrange horn should have a length not exceeding 340 mm and should be free of flare discontinuities.

0 INTRODUCTION

Horn loudspeakers, particularly those used in systems where high-quality sound reproduction is required, have been the subject of debate for a number of years. The use of horns for public-address applications, where high electroacoustic efficiency and good directivity control are of paramount importance, is almost universally accepted as good practice. In fact in many cases there are no alternatives. However, where it is necessary for these useful horn properties to take second place to high sound quality as, for example, in studio monitoring systems, the use of horn loudspeakers is questionable. The controversy about the use of horns in these systems is generally confined to the midrange of frequencies (500 Hz to 10 kHz). The use of horn loudspeakers for the reproduction of low frequencies is often totally impractical due to sheer size limitations. For very high frequencies, horns seem to be acceptable to most people.

Modern recording studio practice can place high demands on a monitor loudspeaker system. Music is increasingly being created using electronic equipment, with the result that many musicians prefer to "perform" in the control room rather than in the studio, using the monitor loudspeakers as extensions of their instruments. During multitrack recording sessions, instruments such as drums need to be reproduced individually at high levels to enable spurious noise from "rattles" or "squeaks" to be detected. Also, many modern studio control rooms are very large and acoustically "dead," requiring a large amount of acoustic power to generate the high sound levels required.

Although these power requirements can be fulfilled by modern midrange direct-radiating loudspeakers, these loudspeakers are frequently working close to their maximum power-handling capability and drive-unit failure has been relatively commonplace. Wasted recording studio time while loudspeaker diaphragms are being replaced can be expensive.

The arguments for the use of midrange horn loudspeakers in studio monitor loudspeakers center around their high electroacoustic efficiency. For the same electric input power, a typical horn–compression-driver combination is capable of delivering 10 times the sound power output of a typical direct-radiating drive unit. Under similar operating conditions, a horn loudspeaker is therefore receiving one tenth of the electric input power of a direct radiating drive unit and, given similar

* Manuscript 1993 December 8; revised 1995 February 21 and November 1.

power-handling capability, is likely to prove more robust and reliable.

The superior power output capabilities of studio monitoring systems containing horn loudspeakers compared to those with direct-radiating drive units are seldom disputed. However, opinions on the perceived sound quality of horns are generally polarized on two extremes. On the one hand there are people for whom horn loudspeakers are capable of the very highest sound quality with an "immediacy" or "clarity" unobtainable with other systems. On the other hand there are people who actively dislike the reproduction of sound over horn loudspeakers and claim that horns have a characteristic sound, often described as "honky" or "quacky," which allows a horn to be identified as a horn. The physical causes of this characteristic horn sound have not (to the authors' knowledge) been documented previously. The identification of such could result in the design of midrange horns for studio monitors that combine the high power output of current designs with the neutral sound quality of the best direct radiators. To provide possible clues as to the physical causes of any characteristic horn sound, a blind listening test was set up allowing listeners to compare horns with other horns and with direct-radiating loudspeakers directly.

The work contained herein was carried out as part of a 3-year research project on midrange horns for studio monitor loudspeakers [1]. Other work on the modeling and measurement of horn behavior is reported in [2].

1 DESCRIPTION OF LISTENING TEST

1.1 Test Objectives

A wide selection of horn loudspeakers designed to operate in the midfrequency range was made available along with a selection of equivalent direct-radiating loudspeakers. As most of these units were not capable of reproducing the entire audio frequency range, it was decided that the test could not aim at comparing the absolute sound quality of the loudspeakers without "filling in" the rest of the frequency range. It was thought impractical to attempt to interface, both physically and electroacoustically, such a wide variety of midrange units with the rest of a loudspeaker system, and as a result, the test was limited to establishing differences or similarities between the loudspeakers in the midfrequency range only. The test was not designed as an experiment in listener psychology, but as an attempt to use the perception of listeners as a tool for finding the solution to an engineering problem. A blind listening test was therefore designed to answer three questions:

1) Do horns sound different from each other?

2) Do horns sound different from direct radiators?

3) If the answers to 1) and 2) are affirmative, is the difference between horns and direct radiators in general greater than the difference between different horns?

1.2 Experimental Design

Ideally, the listening test would involve many different loudspeakers, many different test signals, and a large number of listeners. Each listener would then compare the reproduction of each signal by each loudspeaker with that of every other loudspeaker, with the order of presentation arranged so as to minimize bias (using a "latin square" arrangement, for example). If this had been attempted, however, the test would have taken a prohibitive amount of time. Thus some compromise in experimental technique was necessary.

The physical properties of the loudspeakers that were likely to be responsible for any "horn sound" were largely unknown, so it was important that as many different midrange horn loudspeakers as possible be included in the test along with some direct-radiating loudspeakers. A large number of test loudspeakers was therefore unavoidable. Also largely unknown was whether any particular type of signal would emphasize the differences or similarities between the loudspeakers, so it was considered important to use a number of significantly different test signals. Finally, in order that the results of the test should be statistically significant, a reasonable number of listeners was required. With these requirements in mind, it was decided that each test loudspeaker would be compared with only four reference loudspeakers chosen as typical examples of their kind, thereby allowing a reasonable number of loudspeakers, signals, and listeners to be tested within practical time limitations.

Many parameters can affect the perceived sound quality of a loudspeaker. In order to extract meaningful information from the results of the test, the influence of a large number of these had to be suppressed. A number of measures were taken to effect this simplification.

1) The effect of a loudspeaker's directivity characteristics on the perceived sound quality is very dependent on the acoustic environment in which listening takes place. To remove the effects of directivity, the test was set up in the large anechoic chamber at the Institute of Sound and Vibration Research. To carry out the test in any "ordinary" room would have introduced a complicated and unquantifiable additional variable into the results. Also, unlike horns intended for use in public-address applications, the control of directivity is not a main requirement for horns intended for studio monitoring. The recording engineers are almost always in the optimum listening position (usually on axis) and only hear the off-axis radiation via reflections which, in the relatively "dead" acoustic environment of a modern recording studio control room, are minimal. Anechoic conditions therefore represent an approximation to the actual acoustic environment under which the monitors would be used. Should this test have failed to reveal the causes of a characteristic horn sound, it would have been desirable to repeat it under controlled semireverberant conditions and *then* introduce directional characteristics as a possible cause.

2) It is well known that the level at which a sound is reproduced strongly affects the perceived sound quality. The sensitivities of the loudspeakers under test covered a range of about 25 dB. Thus a means of adjusting the amplifier gain individually for each loudspeaker had to

be incorporated in the setup. At first it appeared that the gain for each loudspeaker would have to be set for each different sound, since two loudspeakers adjusted to reproduce the same level on one test signal would not reproduce the same level on another due to different frequency response characteristics. While it would be logistically very demanding, the matching of levels for individual sounds was considered undesirable since the reproduction of different sounds at different levels can be considered to be an important part of the characteristic sound of a loudspeaker. To attempt to remove these differences between the loudspeakers would therefore bias the test results. The gain settings were therefore adjusted such that each loudspeaker reproduced band-limited pink noise at the same sound pressure level at the listening position.

3) To ensure, as far as possible, that each loudspeaker reproduced the same bandwidth, fourth-order (24 dB per octave) filters at 1 kHz (high pass) and 6 kHz (low pass) were introduced into the signal path.

4) Finally an acoustically transparent but visually opaque curtain was erected between the loudspeakers and the listener to eliminate the "I see, therefore I hear" phenomenon.

1.3 Test Setup

Fig. 1 shows the listening test arrangement. To ensure that every listener sat on axis to every loudspeaker, only one listener took the test at a time. This also removed the temptation for a listener's comments or actions to influence the judgment of another. The loudspeakers were arranged in a circular arc of 3-m radius subtending 60° around a swivel chair. As sitting in an anechoic chamber can be disconcerting to some people, and for safety reasons, the experimenter was present in the chamber at all times behind a plastic foam wall, which permitted verbal and eye contact with the listener.

1.4 Loudspeakers under Test

A total of 20 loudspeakers were made available for the test. These included 14 compression-driver–horn combinations, four direct-radiating cone loudspeakers, a dual concentric horn, and an electrostatic loudspeaker. The electrostatic loudspeaker, a direct-radiating cone, a compression-driver–horn combination, and the dual concentric horn were chosen as the four references and were labeled A to D, respectively. At any one time, five loudspeakers were set up for listening, and were arranged behind the curtain in the following order from left to right: A, B, sample, C, and D. The term "sample" herein refers to the particular loudspeaker under test. A list of all of the loudspeakers tested, including the references, is contained in Appendix 1.

Among the 16 test loudspeakers were two "experiment controls"—loudspeakers for which the results were thought to be known prior to the test. These were included to give an indication of the accuracy of the test result data. One of the controls, sample 6, was a direct-radiating cone loudspeaker which was nominally identical to reference B; the expected result was therefore 100% similarity with reference B. The other control, sample 3, was a direct-radiating cone loudspeaker which was designed to operate over a lower frequency range than that of the test; the expected result for this control was no similarity with any of the references. The results for the rest of the samples were expected to lie somewhere between these "similar" and "nonsimilar" controls.

1.5 Test Signals

To prevent the listeners from being distracted by any preferences for the sound quality of any particular loudspeakers, it was considered important that the signals used in the test contain as little information content as possible. This ruled out speech and music, which would probably have been the most relevant signals. But many of the transient and steady-state features of these were represented by nine signals—two synthetic and seven recorded "live" sounds. One second of each signal was sampled and repeatedly recorded onto digital audio tape for 3 minutes with short gaps between reach repetition. Two repetitions were played through each loudspeaker in the following order: sample, A, sample, B, sample, C, sample, D, sample, A, . . . until the listener had made a decision. At any time the listener could request that any comparison be repeated or omitted via verbal communication with the experimenter. A list of the signals used, along with brief descriptions and replay levels, is contained in Appendix 2.

1.6 Test Equipment

Throughout the setting up of the test, great care was taken to ensure that conditions were as nearly identical for all loudspeakers and all listeners as was practical. The leads running from the switch box to the loudspeakers were all of the same length, and were rated at 15 A. The switch box itself was tailor-made for the test and was split into two sections—one for the line-level signals with six separate gain controls linked to a six-way switch and input and output buffers, and the other for the loudspeaker-level signals with a six-way switch, rated at 10 A, which was ganged to, but physically separated from, the low-level switch. A preamplifier coupled to a 150-W power amplifier was used to drive the loudspeakers and provide overall gain control. A digital tape recorder (DAT) was used as the signal source, providing rapid indexing to the start of each signal. Care was taken to minimize noise in the setup by avoiding earth loops and ensuring the presence of adequate signal levels at all stages.

1.7 Listeners

In all, 20 persons were kind enough to act as listeners. They included mostly people from the professional audio industry, along with some lay people and some academics, with ages ranging from late teens to middle age. Since the horn debate has risen largely out of the varied opinions of recording studio personnel over the years, it was thought that the experience of some "professional listeners" could be valuable in the search for the characteristic horn sound. The listeners were not screened for hearing abilities.

On average, a complete test took about 4 hours. As quite a high degree of concentration was required, listeners were given frequent opportunities for breaks. Many of the listeners could not afford the time to complete a whole test, so those loudspeakers that they missed were covered by another person.

1.8 Questionnaire

The questionnaire that the listeners were asked to fill out consisted of five columns, marked A, B, C, D, and NONE, and for each sample loudspeaker there were nine rows, marked 1–9. Four such sheets comprise a complete questionnaire to cover the 16 sample loudspeakers. For each sample loudspeaker, and for each of the nine numbered sounds, the listener was asked to tick the column for the reference loudspeaker A, B, C, or D that sounded most similar to the sample loudspeaker. More than one column could be ticked, and the NONE column could be ticked when none of the references were judged to sound similar to the sample. The interpretation of the word "similar" was left up to the individual listener. Minimal briefing was given to the listeners to avoid any preconditioning of a listener's opinions. (An interesting result of the test was finding out what people mean when they say "different" or "similar" in the context of this test and to what extent people differ in this respect.) To have been too prescriptive in this aspect of the test could have led the listeners down a restrictive path, and the elusive horn sound could have been missed.

The tick boxes were large enough for short comments to be written about a comparison if the listener desired; indeed, such additional information was welcomed. Each listener, even those who were finishing another person's test, was given a clean questionnaire so that at no time did any listener have access to a previous listener's comments or results.

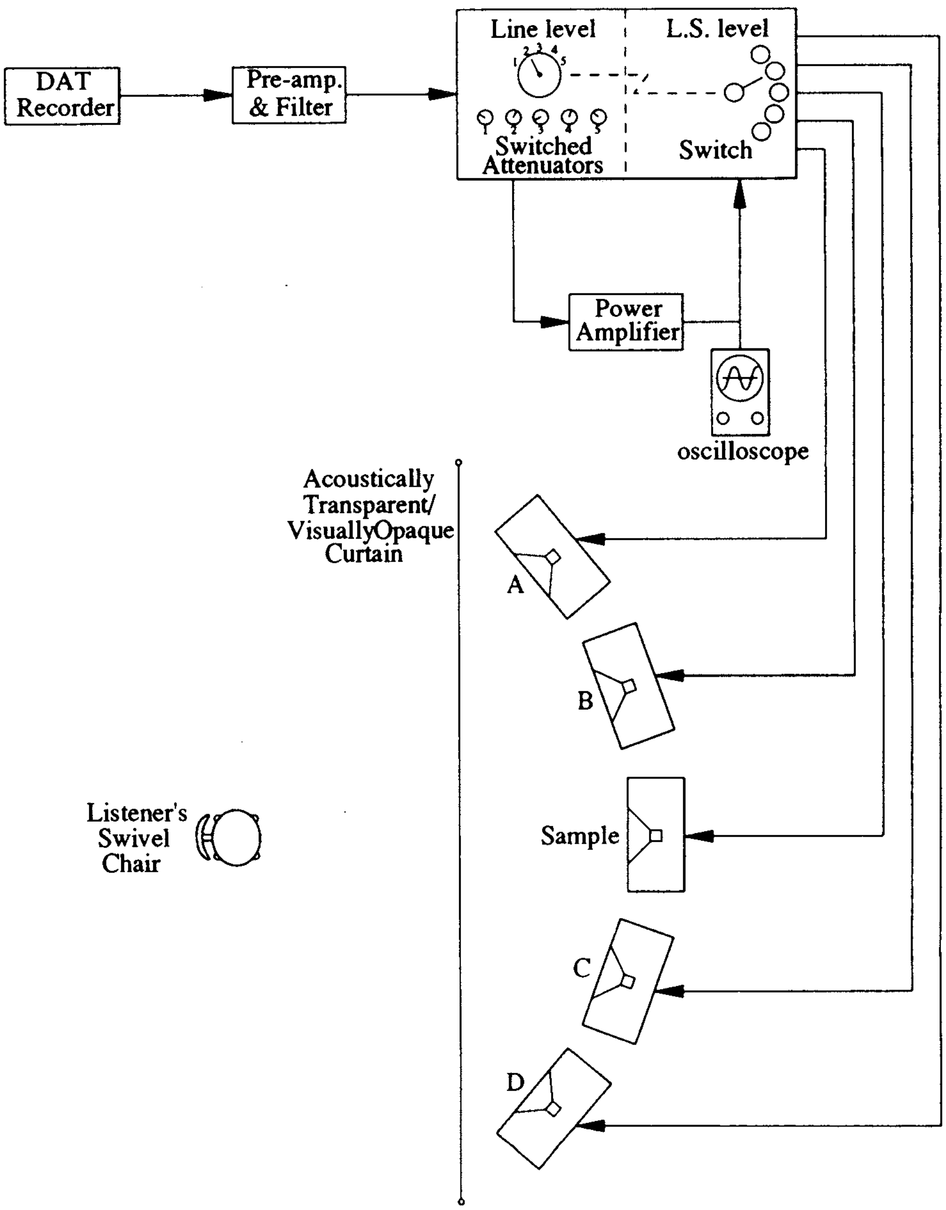


Fig. 1. Listening test setup in anechoic chamber.

2 LISTENING TEST RESULTS

Twelve complete tests were carried out by 20 listeners. Each test contained information on 64 comparisons for nine sounds, a total of nearly 7000 comparisons. Clearly, to present all of these data separately within this paper would be impractical. So for reasons of neatness a breakdown of the data will be presented. For the purpose of analysis, the test results were displayed in three tables showing the number of ticks entered in each questionnaire box. First, results for which ticks were entered in more than one column or those with comments showing reservations alongside them were weighted with half a mark, with unambiguous ticks carrying a whole mark. Second, only the unambiguous ticks were counted. Third, all ticks were given equal weighting. Table 1 shows the unambiguous ticks from all 12 tests for each comparison. The various comments made by listeners about the comparisons will not be itemized in the results, but will be discussed along with the data analysis. The entire set of results can be made available to any interested parties.

3 ANALYSIS OF RESULTS

The analysis of the listening test data was divided into three parts. First, the data were studied in their "raw" form, with conclusions drawn from the consideration of the number of ticks entered for each sample loudspeaker and for each sound. Second, an analysis was performed on measurements taken of the frequency response functions (FRFs) of the sample loudspeakers and on recordings of the test. The results of these analyses were

Table 1. Listening test results—unambiguous ticks.

	Sample 1					Sample 2					Sample 3					Sample 4				
Signal	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None
1	-	1	2	4	1	-	2	4	1	-	-	3	1	1	-	-	6	-	2	1
2	-	2	1	3	3	-	4	2	1	2	-	-	1	2	3	-	6	-	-	2
3	-	3	-	4	1	-	7	1	-	1	1	-	1	1	4	-	6	-	2	1
4	-	2	-	1	4	-	3	-	-	5	-	-	-	-	8	-	5	-	-	4
5	-	1	1	2	2	1	4	-	-	4	-	-	1	-	9	-	6	-	-	4
6	-	5	-	3	3	-	4	-	1	2	-	1	1	1	5	-	3	-	-	3
7	-	1	-	2	4	-	3	1	-	2	-	-	-	-	9	-	7	-	-	2
8	-	4	-	5	-	-	5	-	1	-	-	2	-	2	4	-	6	-	1	2
9	-	3	-	3	3	-	9	-	1	-	-	-	-	-	5	-	1	-	2	4

	Sample 5					Sample 6					Sample 7					Sample 8				
Signal	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None
1	-	3	2	2	-	-	6	-	-	-	-	5	1	-	-	-	1	-	4	1
2	-	5	1	1	1	-	7	-	-	3	-	5	2	-	2	-	1	-	5	3
3	-	2	1	4	1	-	8	-	-	-	-	3	2	-	1	-	1	1	3	1
4	-	1	-	1	4	-	11	-	-	1	1	2	-	1	5	-	1	1	2	4
5	-	2	1	2	4	-	12	-	-	-	-	1	-	1	4	-	1	-	2	6
6	-	2	1	2	1	-	8	-	-	1	-	3	2	-	1	-	2	1	1	3
7	-	-	-	2	6	-	10	-	-	-	-	-	-	-	4	-	-	-	2	8
8	-	6	-	1	2	-	9	-	-	1	1	1	2	-	-	-	1	1	1	3
9	-	3	-	3	3	-	11	-	-	-	-	4	1	1	3	-	2	-	3	4

	Sample 9					Sample 10					Sample 11					Sample 12				
Signal	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None
1	-	3	2	-	-	-	9	-	-	-	-	3	1	-	2	-	3	2	1	-
2	-	3	2	-	2	-	7	-	-	1	1	1	3	-	2	-	1	2	3	2
3	-	-	4	-	-	-	3	2	-	1	-	3	2	-	-	-	3	5	-	-
4	-	-	-	1	8	-	2	-	-	5	-	-	-	-	6	-	-	1	1	5
5	-	1	-	-	9	-	4	-	-	3	-	1	-	-	2	1	-	1	-	6
6	-	2	-	-	3	-	4	1	1	2	-	4	1	-	1	-	2	4	-	3
7	-	-	-	1	5	-	5	-	1	3	-	-	-	-	4	-	-	3	-	4
8	-	2	-	1	3	-	4	-	1	1	-	4	-	1	3	1	2	5	-	1
9	-	-	4	2	2	-	3	-	4	1	1	6	-	-	1	-	-	5	2	1

	Sample 13					Sample 14					Sample 15					Sample 16				
Signal	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None	A	B	C	D	None
1	-	4	-	1	2	-	2	-	3	1	-	6	1	1	-	2	1	4	-	1
2	1	5	2	-	1	-	3	1	1	1	-	4	-	-	2	-	3	-	-	-
3	1	1	2	-	1	1	1	3	1	1	1	2	3	-	2	1	-	6	-	2
4	-	-	1	-	7	-	-	-	-	5	1	-	-	1	6	2	-	-	-	6
5	-	-	-	-	7	1	-	1	-	5	-	-	1	1	5	-	1	1	-	7
6	3	2	3	-	2	1	1	1	-	5	-	1	1	-	7	-	3	4	-	-
7	2	-	1	-	4	1	1	-	-	5	2	-	-	1	5	-	1	-	-	7
8	3	-	4	-	2	1	-	6	-	1	1	1	3	-	4	2	-	5	-	2
9	2	-	3	1	3	-	-	6	-	3	1	1	2	-	3	-	2	4	-	2

then compared to the results from the listening test data analyses. Finally, further analyses of the FRFs and listening test data were performed to attempt to correlate the listening test results with the physical properties of the loudspeaker samples. The first analysis was carried out independently by two of the authors and discussion was deferred until conclusions had been reached. This was thought desirable as any conclusions that were common to both analyses could then be considered more objective and less the result of preconceptions or bias.

The results for the similar and nonsimilar controls were studied first to determine their effectiveness as controls. Table 1 shows the similar control (sample 6) to have 93% of the unambiguous ticks entered in the B column. This result approximates the expected result of 100% similarity with reference B and indicates that the similar control was effective. The results for the nonsimilar control (sample 3) show 72% of the unambiguous ticks to be entered in the NONE column, with the rest of the ticks distributed fairly evenly among the four references. This result indicates that listeners, on the whole, were undecided as to whether sample 3 was similar or not to any one of the references. There is clearly a strong bias toward the expected result of no similarity with any reference. The nonsimilar control appears to have been fairly effective, although not as effective as the similar control.

The increased spread in results for the nonsimilar control compared to the similar control can be explained by considering the concepts of similar and nonsimilar. The confidence in a result indicating a similarity is likely to be higher than one indicating no similarity, due to the precision inherent in the two terms. To be exactly similar yields a precise description of the comparison, whereas to be exactly dissimilar does not. It may be concluded, therefore, that the choice of controls was valid for the experiment, and that any nearly unanimous positive results can be believed.

3.1 Analysis of Listening Test Recordings and Measurements

At the end of the listening test, each signal was played through each of the loudspeakers in turn and recorded onto digital tape via a measurement microphone placed at the average position of a listener's head. The recorded time waveforms were then transferred onto a computer along with the FRFs of each of the loudspeakers measured using random noise and dual-channel fast Fourier transform (FFT) techniques.

A direct comparison between the time waveforms for different samples and references proved difficult and was restricted to qualitative, eyeball assessment only. An important property of a signal, and one for which comparisons are possible, is its frequency spectrum. The spectra for each of the signals reproduced by the loudspeakers were calculated using FFT techniques and stored for comparison with other spectra. The spectra of each signal, reproduced by each sample, were compared to the equivalent spectra reproduced by the four reference loudspeakers using the mean-squared-error technique described in Appendix 3. The result of these comparisons is shown in Table 2. The numbers shown have arbitrary units, with a high number representing a close similarity and a low number a large difference. The numbers highlighted in bold and underlined are the highest values of spectral similarity for a particular signal. Note that in these results, the NONE column has been replaced by an ABSOLUTE column, which refers to comparison with the source signal.

3.2 Further Analysis of Listening Test Recordings and Measurements

In an attempt to find relationships between the test results and the physical characteristics of the sample loudspeakers, the measured FRFs for the samples were analyzed in detail.

A form of power cepstrum, derived from the FRFs of each of the reference and sample loudspeakers, was also used in the analysis. The cepstra were calculated from the log amplitude of the FRFs after deconvolution of the bandpass filter and subsequent corrective filtering. Details of the method used are contained in Appendix 4 and [3]. These power cepstra are useful for separating out parts of a signal (or response) that are time-separated from the main signal (or response), such as reflections or echos. The presence of such a reflection in a response would appear as comb filtering on an FRF plot, which is difficult to detect reliably. The same reflection would appear as a clear "spike" displaced along the x axis (units of time) on a cepstrum plot, allowing it to be easily identified. A study of the power cepstra of the loudspeakers can be useful for finding the physical cause of some of the response irregularities that they may possess.

4 DISCUSSION OF RESULTS

The results of the experimental controls, samples 6 and 3, are discussed in detail in Section 3. These controls, as well as helping to validate the test results, were expected to define a range of results from similar to nonsimilar, over which the results for the rest of the samples would lie. A scale of similarity can be set up between the two control extremes such that results close to those for control sample 6 can be considered similar and results close to control sample 3, nonsimilar. Using this scale, some observations can be made from a study of the test results.

One obvious result is that more ticks have been entered in the reference B column than in any other column. Reference B has been used as the midrange driver in some popular studio monitoring systems. Its choice for use in these systems may be due to the broad similarity to a wide range of other loudspeakers evident from these results.

Other clearly defined results indicate that samples 2, 4, and 10 are similar to reference B; samples 1 and 5 are fairly similar to both references B and D; sample 12 is similar to reference C; sample 8 is quite similar to reference D; sample 9, like the nonsimilar control, is similar to none of the references; samples 7, 11, 13, 14, 15, and 16 all show some similarity with references B, C, and D, with sample 13 also showing some similarity with reference A, and sample 16 more similarity with reference C.

Considering the test signals individually, it is clear that signals 4, 5, and 7 (see Appendix 2) gave rise to many ticks being entered in the NONE column. These signals are all wide-band noise-type signals, and this result seems to indicate that this type of signal is the most critical for showing differences between loudspeakers. The consistency in the strongly similar results, such as those for the similar control, across the range of test signals indicates that conclusions drawn from these positive results would very probably remain consistent, even if extended to a wide range of program material and to subjective similarity in general.

4.1 Comparison between Listening Test Results and Measurements

The objective of this comparison was to discover whether the listening test results can be explained in terms of the similarity, or otherwise, between the frequency spectra of the reproduced test signals. This will be the case if the characteristic sound of the loudspeaker samples in the test is due entirely to on-axis frequency response and does not depend on phase response or nonlinearity. (The test was conducted under anechoic conditions, so directivity should not be a factor.) Table 2 shows the waveform spectral similarity for each of the samples compared to each of the references, for each signal. The numbers shown have arbitrary units and have

Table 2. Waveform spectral similarity (arbitrary units).

	Sample 1					Sample 2					Sample 3					Sample 4				
Signal	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.
1	56	**125**	36	46	93	**82**	72	26	77	93	**76**	54	21	69	62	37	**51**	32	31	46
2	45	**81**	51	41	95	34	**42**	31	25	50	22	26	21	**40**	24	33	**39**	27	24	40
3	16	24	23	**28**	23	24	**36**	32	28	42	**38**	32	29	25	49	19	**35**	25	30	28
4	24	**44**	37	27	43	17	**30**	20	18	30	17	16	18	**21**	18	18	**39**	28	22	35
5	22	**42**	32	26	37	20	**34**	24	19	36	21	22	21	**26**	26	17	**36**	26	21	29
6	26	**46**	36	25	45	22	**39**	30	22	41	22	25	24	**29**	31	18	**37**	28	22	29
7	24	**47**	38	26	46	18	**33**	23	18	33	18	19	19	**25**	21	17	**38**	29	20	32
8	32	**61**	37	26	54	31	**43**	31	24	52	21	26	21	**31**	27	25	**41**	30	23	36
9	18	**34**	30	26	29	19	**33**	23	19	35	**24**	21	20	22	28	16	**34**	27	21	27

	Sample 5					Sample 6					Sample 7					Sample 8				
Signal	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.
1	49	**94**	38	46	78	50	**97**	37	42	69	63	**119**	32	56	105	75	**101**	31	91	159
2	41	**67**	49	35	86	45	**94**	47	38	104	41	**61**	40	40	69	50	**71**	48	55	64
3	16	23	23	**25**	24	23	**67**	27	36	40	32	**40**	**40**	26	69	18	26	27	**36**	27
4	24	**44**	37	26	51	24	**68**	31	31	54	26	**42**	34	27	52	37	32	30	**39**	35
5	22	**41**	35	24	43	21	**66**	32	30	47	25	**40**	38	23	53	28	**35**	31	32	37
6	25	**44**	40	24	48	23	**71**	36	32	49	26	**41**	**41**	23	54	34	**43**	32	32	45
7	23	**46**	39	24	50	23	**72**	34	30	53	26	**43**	36	25	52	**37**	36	31	**37**	39
8	33	**52**	42	23	60	27	**66**	32	32	50	33	**41**	39	23	57	38	**56**	36	31	61
9	18	**34**	**34**	24	32	20	**62**	29	30	45	22	37	**39**	23	44	21	29	**32**	31	29

	Sample 9					Sample 10					Sample 11					Sample 12				
Signal	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.
1	44	**54**	42	53	55	62	**110**	29	42	85	48	**67**	28	34	61	56	**81**	31	39	73
2	**46**	45	43	29	52	50	**70**	38	37	80	**41**	40	37	23	55	**63**	54	43	30	65
3	17	20	**24**	21	23	26	**39**	36	30	50	28	**37**	36	25	58	35	26	**65**	23	55
4	27	27	**35**	20	31	25	**44**	36	30	50	25	**43**	32	24	62	28	33	**39**	24	37
5	24	25	**30**	19	28	23	**44**	36	28	50	28	**39**	35	21	65	33	31	**43**	21	45
6	26	25	**30**	18	30	25	**47**	39	29	53	30	**39**	**39**	22	66	36	34	**46**	21	55
7	27	27	**34**	20	31	25	**46**	36	29	51	26	**43**	35	23	62	30	34	**41**	23	44
8	32	26	**37**	17	33	30	**48**	34	30	56	**41**	40	40	22	72	**47**	34	43	20	61
9	20	24	**33**	19	26	21	**41**	37	28	45	27	**36**	32	21	65	26	29	**46**	22	38

	Sample 13					Sample 14					Sample 15					Sample 16				
Signal	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.	A	B	C	D	Abs.
1	49	**71**	32	34	60	79	**118**	32	52	127	44	**68**	32	34	53	73	**138**	33	58	132
2	54	**60**	49	32	98	55	**67**	61	37	88	49	**59**	39	42	65	60	**73**	45	38	100
3	34	30	**53**	24	62	22	25	**33**	28	34	39	30	**51**	24	54	37	**38**	35	27	78
4	30	32	**39**	29	40	**36**	32	**36**	31	43	26	29	**35**	27	33	40	**44**	32	33	62
5	33	32	**47**	23	46	32	36	**39**	28	50	29	32	**40**	23	42	32	**42**	33	27	61
6	35	35	**50**	24	55	39	**41**	**41**	30	64	30	35	**43**	24	47	36	**39**	33	27	59
7	31	34	**42**	26	44	**38**	36	**38**	31	51	27	32	**36**	26	38	39	**44**	33	31	63
8	43	35	**49**	22	58	45	**54**	43	28	82	34	**37**	**37**	24	49	35	**48**	32	28	56
9	28	30	**49**	24	39	26	33	**38**	28	41	25	30	**41**	23	36	26	**41**	32	26	54

been scaled to ease interpretation.

The spectral comparison for the similar control, sample 6 versus reference B, shows high numbers, indicating that there is high similarity between the spectra of all signals reproduced via sample 6 and reference B. The comparison for the nonsimilar control, sample 3, shows low numbers, indicating dissimilarity between the spectra of all signals reproduced via sample 3 and any of the references, with the exception of signal 1. The numbers for signal 1 are generally higher than those for the other signals due to the narrow bandwidth of this signal leading to less possibility for error between spectra.

A comparison between the spectral similarity results (Table 2) and the listening test results (Table 1) shows good agreement. Clearly this agreement between the listening test results and the comparisons between the spectra of the signals reproduced by the test loudspeakers indicates that a large part, but not all, of the sound of the test loudspeakers can be described in terms of their on-axis amplitude frequency response. This finding is in accordance with those of other researchers. (For example, see Toole [4] and Gabrielsson et al. [5].)

The key to the disparity between the listening test results and the spectral comparisons may lie in the phase response of the loudspeakers. Fig. 2 shows the measured FRFs for each of the reference loudspeakers (through the band-pass filter). It can be seen that the phase response of reference B is quite different from those of references A, C, and D. Fig. 3 shows the filtered FRF of sample 16, which will be used here as an example to attempt to explain differences between the listening test results and the spectral comparisons. The amplitude of the FRF of sample 16 can be seen to be most similar to that of reference B. This result is borne out by the spectral similarity results that show sample 16 to be most similar to reference B for all nine signals. The listening test results, however, show sample 16 to be similar to reference C for five of the signals, similar to no references for three signals, and similar to reference B for only one signal. Fig. 3 shows that sample 16 has a phase response similar to references C and D. The amplitude response of sample 16 is unlike that of reference D, so the listening test results for sample 16 may possibly be explained using a combination of the amplitude and the phase of the FRFs.

4.2 General Answers to Questions Posed in Section 1.1

The listening test was designed to answer the three questions listed in Section 1.1. In this section the test results are studied in a general manner in an attempt to give broad answers to these questions.

Question 1) Do horns sound different from each other? Thirteen of the 16 sample loudspeakers and two of the

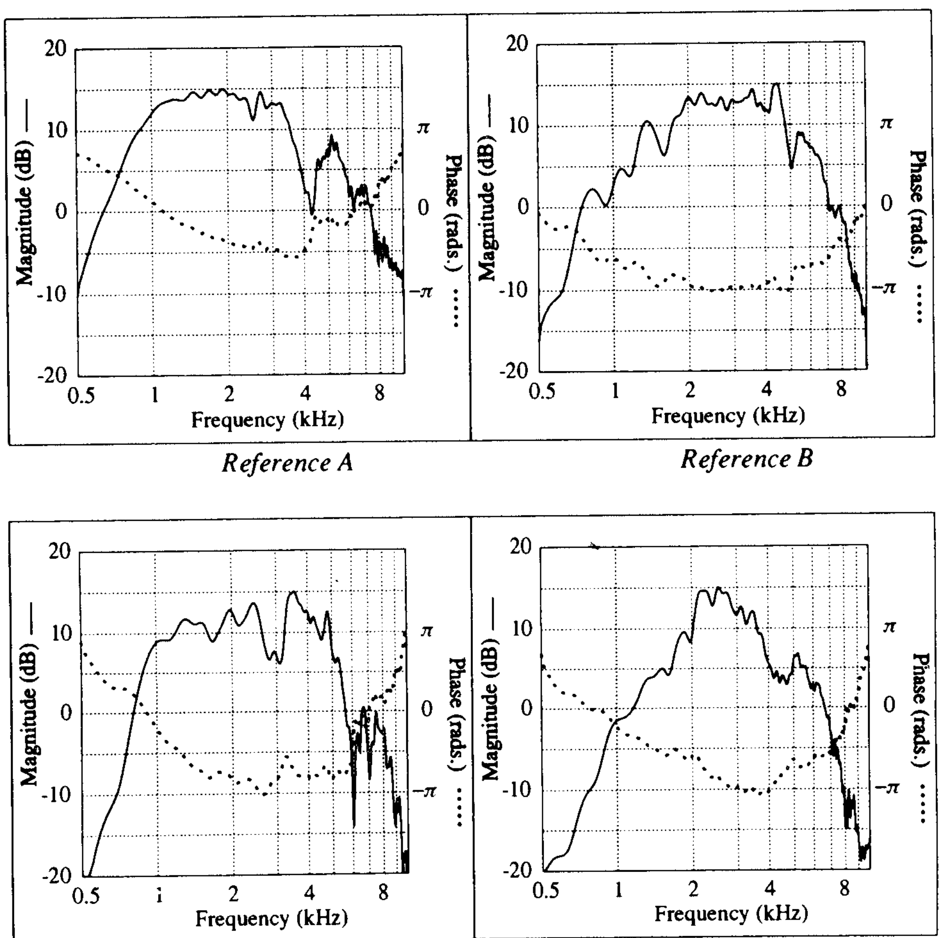


Fig. 2. Frequency response functions of reference loudspeakers.

references are horns, giving a total of 26 horn–horn comparisons for each of the nine signals. The number of unambiguous ticks for all of these comparisons total 272, giving an average of 10.5 ticks per loudspeaker comparison, or an average of 1.16 ticks per signal. If this average is compared with 9.11 ticks per signal for the similar control and 0.53 ticks per signal for the non-similar control, it is clear that in this test, in general the horns do not sound similar to each other. When it is considered that 11 of the 13 horn loudspeaker samples were driven by the same driver, the wide spread in the results for these samples shows that horns do sound significantly different from each other.

Question 2) Do horns sound different from direct radiators? The horn–direct-radiating comparisons that will be considered here do not include the controls. They are the 13 sample horns versus reference B and the two direct-radiating samples versus references C and D, a total of 15 comparisons. The number of unambiguous ticks for these comparisons totals 270, giving an average of 18 ticks per loudspeaker comparison, or two ticks per signal. Again, comparison with the results for the two controls shows that in this test, in general, the horns do not sound very similar to the direct radiators. The results do, however, indicate that some similarity exists between the sound of some of the horns and that of the direct radiators (particularly reference B).

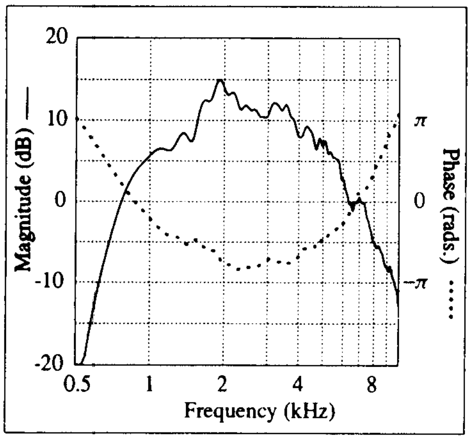


Fig. 3. Frequency response function of sample 16.

Question 3) Is the difference between horns and direct radiators greater than the difference between horns? A comparison between the answers to questions 1) and 2) shows that there is a greater difference between horns than there is between horns and direct radiators. However, it must be borne in mind that most of the similar results mentioned in the answer to question 2) were for comparison between the samples and reference B. As mentioned, the reproduction of sound through reference B appears to be more representative than that of the other three references of the reproduction through a wide variety of loudspeakers, so this result is perhaps not surprising. It can be concluded though that *no* evidence exists from the results of this listening test to show that horns in general sound more different from direct-radiating loudspeakers than they do from each other, and that some horns do sound similar to some direct radiators.

4.3 General Discussion of Results

It was found that, in general, the similarity between the horn loudspeakers in the test and the direct-radiating loudspeakers was at least as strong as the similarity between different horns driven by similar, or even the same, drivers. The horn samples appear to polarize more or less into two different sounds, some having a strong similarity with the direct-radiating reference B and little or no similarity with the horn reference C, and others having a similarity to reference C and little similarity with reference B. Table 3 shows the horn samples grouped according to which reference they sound most similar to. Those samples with asterisks showed particularly strong similarity to the reference. All of the horn samples were fitted with the same Emilar driver, with the exception of sample 11, which was fitted with a similar Emilar driver, and sample 16, which was fitted with a JBL driver. Reference C was fitted with a third Emilar driver.

Table 3. Horn loudspeaker samples grouped according to similarity.

Sample	Manufacturer/Type	Flare Material	Flare Rate	Length (mm)	Mouth Size
	Horns with similarity to reference B (Son Audax direct radiator)				
1	Vitavox exponential	Aluminum	Medium	340	Medium
4	AX1 axisymmetric*	Glass-fiber	Low	230	Small
5	Reflexion Arts	Glass-fiber	Medium	330	Medium
7	Reflexion Arts, no lips	Glass-fiber	Medium	240	Medium
10	Fostex sectoral*	Wood	High	440	Large
11	JBL axisymmetric	Aluminum	Low	250	Small
	Horns with similarity to reference C (Fostex doctoral)				
C	Fostex sectoral	Alumimum	Medium	500	Large
12	Altec sectoral*	Aluminum	Medium	530	Large
13	Altec multicellular	Aluminum	Low/med	600	Large
14	Starr gramophone	Wood	Low	650	Medium
15	Vitavox sectoral	Aluminum	Medium	450	Large
16	JBL biradial*	Composite	Medium	400	Medium
		Others			
8	AX2 axisymmetric	Glass-fiber	High	230	Medium
9	Yamaha sectoral	Aluminum	Medium	350	Medium

Alongside each sample in Table 3 are the materials from which the flares of the horns are constructed, the flare rate of the horn, the length of the flare, and the horn mouth size. It is clear from this grouping that horn length plays an important part in deciding which of the two references a particular horn is more similar to. Horns with diaphragm-to-mouth lengths of less than 340 mm sound similar to reference B, and those with lengths greater than 400 mm sound similar to reference C, which is also a long horn. The exceptions to this rule are samples 8, 9, and 10.

Sample 8, the AX2 axisymmetric horn, showed overall similarity with reference D, with little similarity to references A, B, or C. The main difference between this horn and the other short horns is the high flare rate, giving the horn a higher cutoff frequency and also an almost total lack of mouth reflections. Sample 9, the Yamaha sectoral horn, showed some similarity with both references B and C, but measurements indicate a very uneven frequency response and the sample was generally considered by listeners to sound "strange." This horn differs from the other short horns in having an abrupt change in flare rate and cross-sectional shape partway along the flare, giving rise to the response aberrations and probably the "strange" sound. Sample 10, the wooden Fostex sectoral horn, also has a change in cross section and flare rate, but the flare in the horizontal plane after the change is extremely rapid (included angle 140°), leading to a very wide, almost semicircular mouth. It appears as if the sound of this horn were dictated by the short throat section of the horn, the rest of the flare acting more or less as "lips" for controlling directivity.

The polarization in sound between the short horns and the long horns may be explained by considering the time interval between a signal and any changes that may be imparted on the signal by the loudspeaker. Reflections from the mouths of the short horns are radiated about 1–2 ms after the signal has been radiated, and those from the long horns about 2–4 ms after the signal. It appears from the FRF for reference B (Fig. 2) and from the spectral similarity results (Table 2) that the various resonances and reflections in the direct-radiating cone give rise to irregularities in the FRF that are of a similar nature as those due to reflections in the short horns. The mouth reflections in the long horns are generally less severe than those of the short horns as the mouth is larger. This can be seen from the generally smoother FRFs for the longer horns. From the various comments made by the listeners, both orally during the test and on the questionnaires, it is clear that the longer horns can be more reliably identified as horns. Only one short horn, sample 11, was ever identified as a horn, and then only by one "golden eared" professional sound engineer. There are two possible reasons for this. 1) Because of similar response irregularities, the short horns are mistaken for direct-radiating cone loudspeakers, and 2) since the reflections from the long horns occur after a longer delay, they are more audible. These hypotheses agree with the observation that of the two horns in the test that produce negligible mouth reflections, samples 8 and 13, neither was ever identified as a horn, and the short horn, sample 8, did not sound like the direct-radiating reference B. It should be noted, however, that hypothesis 2) is not in agreement with results from the field of room acoustics on the audibility of a single reflection.

Little or no evidence exists from the results of this listening test that horn flare construction material, flare rate, or shape (sectoral, exponential, and so on) have much effect on the on-axis sound of a horn under anechoic or free-field listening conditions. It is expected, however, that under reflective or reverberant listening conditions the directivity properties of a horn, controlled by the shape and size of flare, will affect the perceived sound quality.

5 PRACTICAL IMPLICATIONS OF RESULTS

It appears from the results of this listening test that the "characteristic horn sound" is due to reflections from the mouths of long horns. Many studio monitor loudspeaker systems have used, and continue to use, long horns for the reproduction of the midfrequency range, and in such systems the presence of the horn can be heard by listeners even though the measured frequency response function, both on and off axis, may be comparatively smooth. These long horns have been "borrowed" from public-address technology, for which the large flare is highly desirable to maintain constant-directivity properties. Such a large flare is not required for studio monitoring purposes, where directivity properties more in line with those of direct-radiating loudspeakers are required, and the flare exists simply for its acoustic transformer properties to increase electroacoustic efficiency. Clearly, a horn used for the reproduction of the midfrequency range in a studio monitoring system should have a diaphragm-to-mouth length of less than 340 mm if the characteristic horn sound is to be avoided.

When fitted with a high-quality driver, the FRF (and hence aspects of sound quality) of a horn, at frequencies above cutoff, is determined principally by reflections from any flare discontinuities, including the mouth termination. The results of this test show that the reflections from discontinuities in the flare of the short horns lead to response aberrations which are similar in character to those due to diaphragm problems in conventional direct-radiating loudspeakers. The relatively small rigid diaphragm in a high-quality compression driver does not suffer such problems, however, so a short horn devoid of flare discontinuities should possess frequency response characteristics that are at least the equal of the best direct-radiating drive units.

Sample 8 in the test was a prototype horn designed to possess minimal flare discontinuities. When coupled to the driver used in the tests, the band-pass filtered FRF of the combination, shown in Fig. 4, is seen to be uneven. Fig. 5 shows the power cepstrum for the combination. The presence of spikes at about 0.3-ms intervals shows a strong reflection from a discontinuity approximately 50 mm from the diaphragm. This is the distance from the diaphragm to the flange between the driver and the horn, and close inspection of the combination revealed a slightly different flare rate between the throat section of the driver and that of the horn. The horn was later coupled to a driver having a matching flare rate, and the response aberrations (Fig. 6) and the cepstral spikes (Fig. 7) were seen to disappear (see [6], [7]).

Public-address style constant-directivity horn designs often rely on diffraction edges for directivity control. These edges represent severe flare discontinuities, and so these horns are generally less suitable for studio monitoring than more conventional designs. It is desirable that horns for studio monitoring or other applications where high-quality sound reproduction is required should be devoid of any flare discontinuities, including the mouth termination.

6 CONCLUSIONS

Answers to the three questions posed in Section 1 were concluded from the results and can be summarized as follows.

- Horns do sound different from each other, even when fitted with the same driver.
- Some similarity exists between the sound of the horns and the sound of the direct radiators.
- No evidence exists that horns in general sound more different from direct radiators than they do from each other.

A set of conclusions, based on the listening test results and measurements, have been reached. These conclusions can be summarized as follows.

- The horn samples appear to polarize into two different sounds—strong similarity with reference B, a direct-radiating cone, or with reference C, a sectoral horn.
- The polarization appears to be dependent on the length of the horn with horns of less than 340 mm between diaphragm and horn mouth sounding similar to the direct-radiator reference, and those with more than 400-mm length sounding similar to the horn reference (a "long" horn).
- The various reflections and resonances in the cone of the direct-radiator reference give rise to irregularities in the frequency response function that are similar to those due to the mouth reflections in the short horns.
- The longer horns were more reliably identified as horns by the test listeners than the short horns.
- The two horns having minimal mouth reflections, one long and one short, were not identified as horns and did not sound similar to the direct-radiating reference.

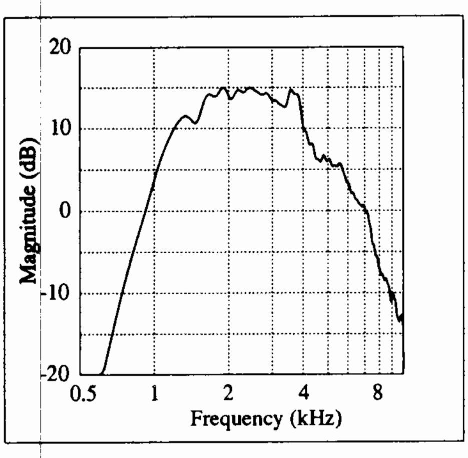


Fig. 4. Magnitude of filtered frequency response function of sample 8.

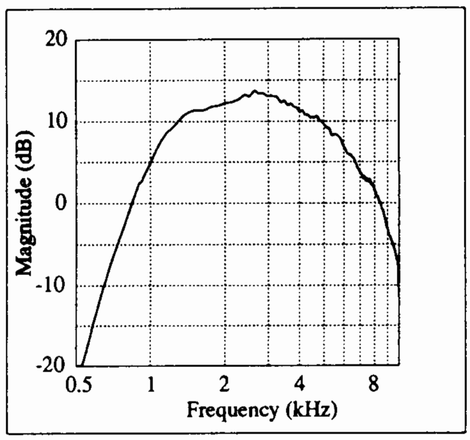


Fig. 6. Magnitude of filtered frequency response function of horn of sample 8 with new driver.

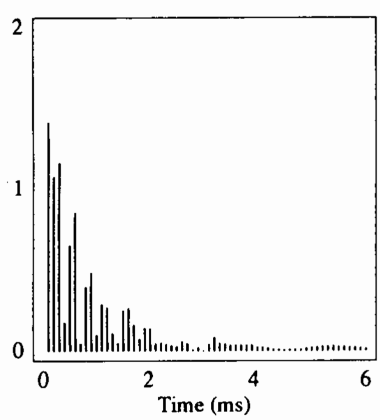


Fig. 5. Power cepstrum of sample 8.

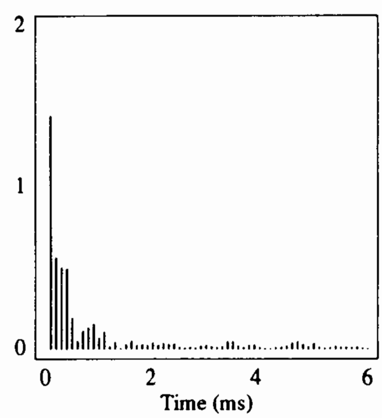


Fig. 7. Power cepstrum of horn of sample 8 with new driver.

- Horns intended for the reproduction of the midfrequency range in studio monitor loudspeakers should have a diaphragm-to-mouth length of less than 340 mm and should be devoid of flare discontinuities.
- Agreement between the listening tests and spectral similarity results indicate that a large part, but not all, of the "sound" of the test loudspeakers under free-field conditions can be described in terms of their on-axis amplitude frequency response.

7 REFERENCES

[1] K. R. Holland,"A Study of the Physical Characteristics of Mid-Range Horn Loudspeakers and Their Relationship to Perceived Sound Quality," Ph.D. thesis, University of Southamption, Southampton (1992).

[2] K. R. Holland, F. J. Fahy, and C. L. Morfey, "Prediction and Measurement of the One-Parameter Behavior of Horns," *J. Audio Eng. Soc.*, vol. 39, pp. 315–337 (1991 May).

[3] K. R. Holland, "The Use of Cepstral Analysis in the Interpretation of Loudspeaker Frequency Response Measurements," *Proc. Inst. Acoust.*, vol. 15, no. 7, Reproduced Sound 9 (1993).

[4] F. E. Toole, "Loudspeaker Measurements and Their Relationships to Listener Preferences: Parts 1 and 2," *J. Audio Eng. Soc.*, vol. 34, pp. 227–235 (1986 Apr.); ibid., pp. 323–348 (1986 May).

[5] A. Gabrielsson, B. Lindstrom, and T. Ove, "Loudspeaker Frequency Response and Perceived Sound Quality, *J. Acoust. Soc. Am.*, vol. 90, no. 2 (1991).

[6] K. R. Holland, F. J. Fahy, and P. R. Newell, "Axi-Symmetric Horns for Studio Monitor Systems," *Proc. Inst. Acoust.*, vol. 12, no. 8, Reproduced Sound 6 (1990).

[7] P. Newell and K. R. Holland, "Do All Mid-Range Horn Loudspeakers Have a Recognizable Characteristic Sound?," *Proc. Inst. Acoust.*, vol. 12, no. 8, Reproduced Sound 6 (1990).

APPENDIX 1
DETAILS OF LOUDSPEAKERS IN LISTENING TEST

Note: Three Emilar EK175 compression drivers were used in the test. These have been labeled nos. 1, 2, and 3.

Reference A: Quad electrostatic loudspeaker (early type). Full-range electrostatic loudspeaker consisting of three separate radiating panel areas for low, mid, and high frequencies.

Reference B: Son Audax PR17/HR100/1AK7. Midrange direct-radiating paper-cone loudspeaker of nominally 6½-in (165-mm) diameter.

Reference C: Fostex H351/HA21 horn/Emilar EK175 driver (no. 3). Large sectoral horn of cast aluminum construction, coupled to a compression driver with an aluminum diaphragm of nominally 2-in (51-mm) diameter and a plastic phase plug.

Reference D: High-frequency horn section of a Tannoy Puma dual-concentric loudspeaker. Axisymmetric horn using paper cone of low-frequency driver as the outer part of the horn. The diaphragm and phase plug of the driver are both aluminum.

Sample 1: Vitavox exponential horn/Emilar EK175 driver (no. 1). Medium-sized exponential horn of cast aluminum construction coupled to a compression driver similar to reference C.

Sample 2: JBL2105. Midrange direct-radiating paper-cone loudspeaker of nominally 5-in (127-mm) diameter.

Sample 3: JBL2121. Lower midrange direct-radiating paper-cone loudspeaker of nominally 10-in (254-mm) diameter. This sample was designed to operate over a lower frequency range than that of the test and was included as a nonsimilar experimental control.

Sample 4: AX1 horn/Emilar EK175 driver (no. 1). Short axisymmetric horn of glass-fiber construction with a low flare rate and small horn mouth. Compression driver as sample 1.

Sample 5: Reflexion Arts horn/Emilar EK175 driver (no. 1). Medium-sized horn constructed of mineral-loaded glass-fiber. Horn flare is rectangular in cross section with a smooth, rapid exponential horizontal flare and a shallow, straight-sided vertical flare. Compression driver as sample 1.

Sample 6: Son Audax PR17/HR100/1AK7. This sample is nominally identical to reference B, originating from the same production batch, and is included as a similar experimental control.

Sample 7: Reflexion Arts horn without lips/Emilar EK175 driver (no. 1). As sample 5, but with the mouth "lips" sawn off flush with the mounting flange.

Sample 8: AX2 horn/Emilar EK175 driver (no. 1). Short axisymmetric horn of glass-fiber construction with a rapid flare rate terminating in a medium-sized mouth. Compression driver as sample 1.

Sample 9: Yamaha horn/Emilar EK175 driver (no. 1). Medium-sized sectoral horn of cast aluminum construction. Compression driver as sample 1.

Sample 10: Fostex H320 horn/Emilar EK175 driver (no. 1). Large sectoral horn of laminated wood construction with near semicircular horizontal flare. Compression driver as sample 1.

Sample 11: JBL2307 horn with JBL2308 slant plate/Emilar EK175 driver (no. 2). Short axisymmetric horn similar to sample 4, but of cast aluminum construction and fitted with horizontal slant plates at the mouth. Compression driver similar to sample 1, but with different mounting arrangements.

Sample 12: Altec sectoral horn/Emilar EK175 driver (no. 1). Large sectoral horn of cast aluminum construction. Compression driver as sample 1.

Sample 13: Altec 806C horn/Emilar EK175 driver (no. 1). Large multicellular horn with eight individual flares of sheet aluminum construction joined to a single throat via a cast aluminum manifold. Compression driver as sample 1.

Sample 14: Starr "singing throat" horn/Emilar EK175 driver (no. 1). Folded phonograph horn of sheet/solid wood construction. Compression driver as sample 1.

Sample 15: Vitavox séctoral horn/Emilar EK175 driver (no. 1). Large sectoral horn of cast aluminum construction. Compression driver as sample 1.

Sample 16: JBL2370 horn/JBL2426 driver. Medium-sized biradial horn of composite plastic construction and flat front. Compression driver has titanium diaphragm and exponential phase plug.

APPENDIX 2
DETAILS OF SIGNALS USED IN LISTENING TEST

Overall test level: $L_{eq} = 71$ dB.

Signal 1: Chirp. Enveloped swept sinusoid. Frequency swept from 2 to 4 kHz with a ½(1 − cos) envelope. Peak level 71 dB SPL.

Signal 2: Tone burst. Ten cycles of 2.5-kHz sinusoid. Peak level 80 dB SPL.

Signal 3: Flute notes. Two flute notes recorded anechoically using a Brüel & Kjær measurement microphone. Peak level 65 dB SPL.

Signal 4: White noise. A 1-s burst of white noise. Peak level 69 dB SPL.

Signal 5: Pink noise. A 1-s burst of pink noise. Peak level 66 dB SPL.

Signal 6: Slamming book. The slamming shut of a heavy book recorded anechoically using a Brüel & Kjær measurement microphone. Peak level 80 dB SPL.

Signal 7: Waterfall. Small stream waterfall recorded using a Knowles electret microphone. Peak level 71 dB SPL.

Signal 8: Impact. The impact of a peach stone on a 25-ft-high square-section steel statue recorded using a Knowles electret microphone. Peak level 76 dB SPL.

Signal 9: Guitar chord. The chord "open *E*," strummed on an acoustic guitar, recorded anechoically using a Brüel & Kjær measurement microphone. Peak level 61 dB SPL.

APPENDIX 3
CALCULATION OF WAVEFORM SPECTRAL SIMILARITY

Two methods were used to obtain the spectra of the reproduced signal waveforms. 1) The spectra of the recorded waveforms were calculated directly using FFT techniques. 2) The source signals were convolved with the measured frequency response functions of the loudspeakers. The two methods yielded very close results, and the latter method is favored because of better signal-to-noise performance only. The spectra obtained contain 512 linearly spaced lines up to a maximum frequency of 10 kHz.

Various different ways of comparing the spectra were tried. These included all combinations of linear $[V(f)]$, power $[V^2(f)]$, and logarithmic (dB) spectra with linear (512 lines) and logarithmic (third, sixth, and twelfth octave bands) frequency spacing. The results from each method were compared to the listening test results and the method that yielded the best agreement, which was linear spectra $[V(f)]$ with linear frequency spacing, was chosen.

In order to calculate the difference between two spectra, the mean levels need to be matched. To achieve this, the total error between the two spectra is calculated thus:

$$C = \frac{\sum_{n=1}^{N} S_1(n)}{\sum_{n=1}^{N} S_2(n)}$$

where $S_1(n)$ and $S_2(n)$ are the linear spectra $[V(f)]$at frequency n and N is the number of frequency points.

One of the spectra is then multiplied by this total error to remove any broad-band level differences. The root-mean-squared error between the two spectra is then calculated,

$$E = \frac{1}{N} \sum_{n=1}^{N} \sqrt{[S_1(n) - S_2(n)]^2} \, .$$

The waveform spectral similarity is calculated from this figure by inverting, normalizing with respect to the average spectral level, and scaling the root-mean-squared error for comparison with the listening test results,

$$\text{waveform spectral similarity} = 10 \times \frac{\frac{1}{N} \sum_{n=1}^{N} S_1(n)}{E} \, .$$

This figure is calculated for each sample, in comparison with each reference and the source signals, for each signal.

APPENDIX 4
CALCULATION OF POWER CEPSTRA OF LISTENING TEST LOUDSPEAKERS

The power cepstrum of an FRF is the Fourier transform of the log of the amplitude of the FRF. The power cepstrum is useful for pinpointing the physical cause of any irregularities in the amplitude response, such as reflections. A reflection can be difficult to identify in the frequency domain, as it shows as comb filtering, but on the cepstral plot such a reflection would show as a single spike displaced along the time axis, which can be more easily identified. The FRFs of the test loudspeakers are all band-limited by the filter, so the power cepstra would be dominated by the low- and high-frequency rolloffs, thus masking any differences between them. To overcome this problem, the filter response is removed from the FRFs by deconvolution. To eliminate the effect of any response irregularities outside the passband of the filter, the resultant log-amplitude responses are normalized to an average level of 0 dB and weighted by the amplitude response of the filter. To illustrate this, Fig. 8 shows how the log-amplitude response of reference B is processed prior to calculation of the power cepstrum. The power cepstra for the listening test loudspeakers are calculated from these resultant amplitude responses using FFT techniques.

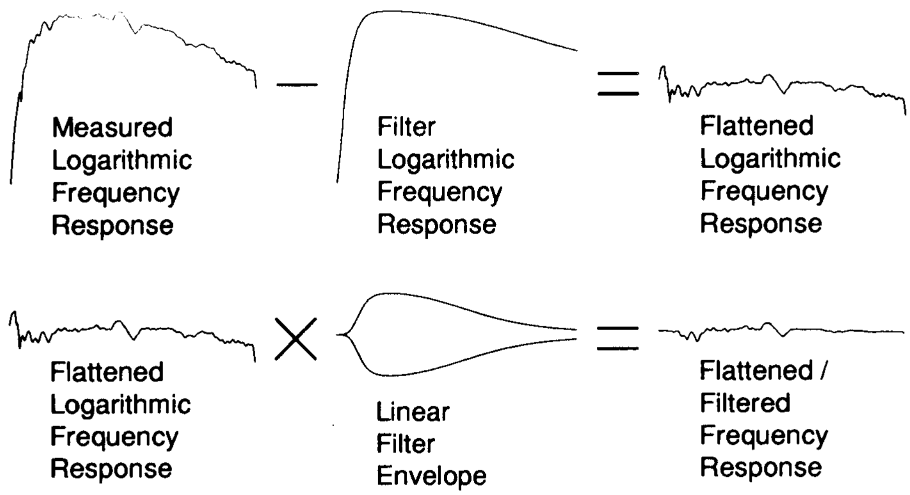
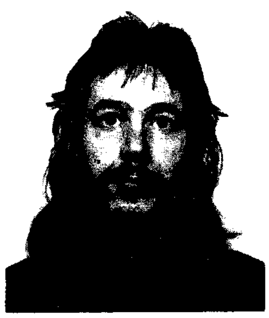
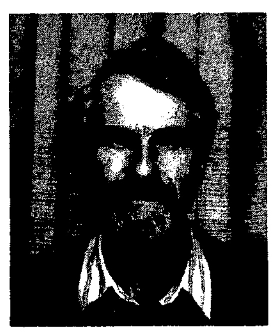


Fig. 8. Diagrammatic representation of processing of loudspeaker frequency response function prior to calculation of power cepstrum.

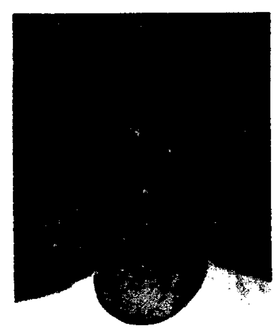
THE AUTHORS



K. R. Holland



F. J. Fahy



P. R. Newell

Keith R. Holland obtained an HND in mechanical engineering in 1978 before working for British Aerospace until 1984. He received a B.Sc. degree in engineering acoustics and vibration in 1987 from the Institute of Sound and Vibration Research, Southampton University, U.K., and a Ph.D. for a thesis on horn loudspeakers in 1993 from the same institute. His other research interests at ISVR include aircraft sound insulation, stereo imaging, vibroacoustics, source location, duct acoustics, vehicle noise, numerical modeling and reciprocity.

Dr. Holland also teaches electroacoustics at under- and post-graduate level at two universities and is a consultant on loudspeaker and studio monitor design. He is a member of the Audio Engineering Society and the Institute of Acoustics.

•

Frank J. Fahy is professor of engineering acoustics and associate director of the Institute of Sound and Vibration Research, Southampton University, U.K. His principal research interests are in the field of vibroacoustics and acoustic measurement. His current activities are concerned with active noise control of low-frequency sound transmission in buildings, uncertainties in vibration prediction based on statistical energy analysis, optimization of truck engine noise shields, electroacoustic horn performance, and applications of reciprocity techniques to the identification of machinery noise sources.

Professor Fahy has written two books: *Sound and Structural Vibration* (1987) and *Sound Intensity* (second edition 1995). He is a member of the Audio Engineering Society and a fellow of the Institute of Acoustics.

•

Philip R. Newell was born in Blackburn, U.K. in 1949. He began working in the sound reinforcement and recording studio industries in 1966, and received his first commission to build a multitrack recording studio, in London, in 1969. In 1970 he moved to Pye Records, working primarily on mobile recordings, and then to the Virgin Records group of companies in late 1971. From 1973 to 1982 he was Technical Director of the recording division of Virgin, during which time he was responsible for the establishment of most of their studios, now owned by EMI.

After an excursion into seaplane flying, he returned to studio and monitor design in late 1984, joining Reflexion Arts for almost three years in the late 1980s, and designing their range of monitors. He has been responsible for the design of over 100 studios in numerous countries, and is a regular contributor to many internationally distributed magazines, many of his contributions being translated and republished in several languages. Throughout his career, he has continued to be actively involved in the recording and production of a wide range of music. In 1991 he moved to Iberia from where he now bases his operations. Philip Newell is the author of a book *Studio Monitoring Design* (1995), and is a member of the Audio Engineering Society.